

# Bell's inequality and extremal nonlocal box from Hardy's test for nonlocality

Sixia Yu

Centre for Quantum Technologies, National University of Singapore, 2 Science Drive 3, Singapore 117542

**Abstract.** Bell showed 50 years ago that quantum theory is nonlocal via his celebrated inequalities, turning the issue of quantum nonlocality from a matter of taste into a matter of test. Years later, Hardy proposed a test for nonlocality without inequality, which is a kind of “something-versus-nothing” argument. Hardy's test for  $n$  particles induces an  $n$ -partite Bell's inequality with two dichotomic local measurements for each observer, which has been shown to be violated by all entangled pure states. Our first result is to show that the Bell-Hardy inequality arising from Hardy's nonlocality test is tight for an arbitrary number of parties, i.e., it defines a facet of the Bell polytope in the given scenario. On the other hand quantum theory is not that nonlocal since it forbids signaling and even not as nonlocal as allowed by non-signaling conditions, i.e., quantum mechanical predictions form a strict subset of the so called non-signaling polytope. In the scenario of each observer measuring two dichotomic observables, Fritz established a duality between the Bell polytope and the non-signaling polytope: tight Bell's inequalities, the facets of the Bell polytope, are in a one-to-one correspondence with extremal non-signaling boxes, the vertices of the non-signaling polytope. Our second result is to provide an alternative and more direct formula for this duality. As an example, the tight Bell-Hardy inequality gives rise to an extremal non-signaling box that serves as a natural multipartite generalization of Popescu-Rohrlich box.

This article is part of a special issue of *Journal of Physics A: Mathematical and Theoretical* devoted to ‘50 years of Bell's theorem’.

Keywords: Bell's inequality, Hardy's test for nonlocality, non-signaling polytope

## 1. Introduction

Fifty years ago, Bell [1, 2] proved that quantum theory is nonlocal by showing that quantum mechanical predictions on a finite set of correlations of local measurements cannot be reproduced by any local realistic model via his celebrated inequality, Bell's inequality. Ever since different types of Bell's inequalities in various scenarios have been proposed and the violations to Bell's inequality have been verified in many experiment situations, finding numerous applications in various quantum informational processes [3] (and references therein). Twenty-five years ago Werner [4] introduced the exact definition of quantum entanglement and proved that entanglement is necessary but not sufficient for the violations of Bell's inequality. However, for pure states these two fundamental features of quantum theory, quantum nonlocality and entanglement, become equivalent, known as Gisin's theorem [5].

The equivalence between the quantum nonlocality and entanglement was established by Gisin [5] and Gisin and Peres [6] in the case of two particles by proving that all entangled two-particle pure states violate a single Bell's inequality, due to Clauser, Horne, Shimony, and Holt (CHSH) [7]. In the case of three qubits Gisin's theorem was proved by Chen *et al.* [8] numerically and by Choudhary *et al.* [9]. In the most general case the equivalence was established by Popescu and Rohrlich (PR) [10] using a set of Bell-CHSH inequalities and more recently by Yu *et al.* [11] using a single Bell's inequality arising from Hardy's test for nonlocality, which is referred to as Bell-Hardy inequality. Years after Bell's theorem, Greenberger, Horne, and Zeilinger (GHZ) [12] provided a compelling argument, which is elegantly formulated by Mermin [13], for the quantum nonlocality of some special states, e.g., GHZ states as well as some graph states [14], without inequality. This kind of "all-versus-nothing" argument is applicable only to three or more particles. Soon after, Hardy [15, 16] provided a nonlocality test for two particles without inequality, which can be regarded as a kind of a "something-versus-nothing" argument.

Mermin [17] formulated a Bell inequality from two-particle Hardy's test, which is equivalent to the Bell-CHSH inequality and is generalized by Cereceda [18] to the case of  $n$  particles. From the point view of the equivalence between nonlocality and entanglement the Bell-Hardy inequality is the most natural generalization of the two-particle Bell-CHSH inequality into many particles. Moreover it is well known that the Bell-CHSH inequality is tight, i.e., a facet of the Bell polytope in the scenario of two observers each of which measures two dichotomic observables. Our first result is to prove that the Bell-Hardy inequality for an arbitrary number of observers is also tight, providing another justification for regarding the Bell-Hardy inequality as a most natural generalization of the Bell-CHSH inequality. As is well known tight Bell's inequalities completely delineate the correlations that are explainable by local realistic theory in a given scenario and thus serve as optimal witnesses of nonlocal correlations.

On the other hand quantum theory is not that nonlocal since it forbids signaling, i.e., sending information via local choices of different measurement settings is impossible.

There are many efforts to understand why this is the case for quantum theory by introducing some additional constraints other than non-signaling, such as information causality [19] and exclusion principle [20] or local orthogonality [21]. In a given scenario, as all local correlations form the so-called Bell polytope, all non-signaling correlations for the so-called non-signaling polytope, which include the quantum theory as a strict subset because there are non-signaling correlations, e.g., the PR box [22], whose correlations are too strong to be reproduced even in quantum theory. An important progress of understanding non-signaling polytope is made by Fritz [23] by establishing a one-to-one correspondence of the extremal points of non-signaling polytope and the facets of Bell polytope in the special scenario of each observer measuring two dichotomic observables. In this case every tight Bell's inequality gives rise to an extremal non-signaling box and vice versa. For example the extremal non-signaling box corresponding to the Bell-CHSH inequality is exactly the well-known PR box. Our second result is to provide an alternative proof for Fritz's duality with a more direct formula. As an example we derive the extremal nonlocal box that corresponds to the tight Bell-Hardy inequality.

## 2. Bell polytope

In an  $(n, s, d)$  scenario there are  $n$  observers that are mutually space-like separated and each observer measures  $s$  observables with  $d$  outcomes. Most generally each observer may perform a different number of measurements and each measurement may have a different number of outcomes. In what follows we consider only the case  $s = d = 2$ , i.e., there are  $n$  observers each of which performing locally two dichotomic measurements.

Let  $n$  observers be labeled with an index set  $I = \{1, 2, \dots, n\}$  and each observer, e.g., observer  $k \in I$ , measures two dichotomic observables  $A_{k|0}$  and  $A_{k|1}$  with two outcomes labeled with  $a_{k|0}, a_{k|1} = 0, 1$  respectively. We denote by  $\mathbf{s} = [s_k]_{k \in I}$  the setting vector with each component  $s_k = 0, 1$  labeling two observables that are measured by the  $k$ -th observer. For a given measurement setting  $\mathbf{s}$  we denote by  $\mathbf{a}_{\mathbf{s}} = [a_{k|s_k}]_{k \in I}$  the corresponding outcome vector with each component  $a_{k|s_k} = 0, 1$  labeling two outcomes of measuring the observable  $A_{k|s_k}$  with  $k \in I$ . Both two vectors  $\mathbf{s}$  and  $\mathbf{a}_{\mathbf{s}}$  are  $n$ -dimensional binary vectors and we denote by  $2^I$  the set of all  $n$ -dimensional binary vectors. Slightly abusing the notations we shall use a binary vector  $\mathbf{a} \in 2^I$  to denote also a subset of  $I$ , i.e., its support  $\{k \in I | a_k \neq 0\}$ , and by  $|\mathbf{a}|$  the number of nonzero components of a vector  $\mathbf{a}$ . For later use we shall build from two vectors  $\mathbf{a}, \mathbf{b} \in 2^I$  a third  $n$ -dimensional binary vector  $\mathbf{a} \wedge \mathbf{b} = [a_k b_k]_{k \in I}$ .

There are in total  $4^n$  events labeled with the setting vector  $\mathbf{s}$  and the corresponding outcome vector  $\mathbf{a}_{\mathbf{s}}$  and each event is associated with a probability  $P(\mathbf{a}_{\mathbf{s}}|\mathbf{s})$  in a probabilistic theory, such as quantum theory or local hidden variable model. In a local realistic model, for a given hidden variable  $\lambda$  distributed according to  $\varrho_\lambda$ , we denote by  $a_{k|s_k}^\lambda \in \{0, 1\}$  the outcome of observer  $k$  measuring observable  $A_{k|s_k}$  for  $k \in I$ . The joint probability of the event that observer  $k$  measures observable  $s_k$  obtaining outcome  $a_{k|s_k}$

with  $k \in I$  then assumes the following local form

$$P(\mathbf{a}_s|\mathbf{s}) = \int d\lambda \varrho_\lambda \prod_{k \in I} \delta(a_{k|s_k}^\lambda, a_{k|s_k}). \quad (1)$$

Every set of  $4^n$  probabilities  $\{P(\mathbf{a}_s|\mathbf{s}) \mid \mathbf{s}, \mathbf{a}_s \in 2^I\}$  of above form, also referred to as a local box here, can be viewed as a vector in a  $4^n$  dimensional space  $\mathcal{R}$ , called as box space. The box space  $\mathcal{R}$  can be conveniently taken as a real subspace of the Hilbert space of  $2n$  qubits and we shall label those  $n$  pairs of qubits by the index set  $I$  with  $\{|\mathbf{s}, \mathbf{a}_s\rangle = |\mathbf{s}\rangle_I \otimes |\mathbf{a}_s\rangle_I \mid \mathbf{s}, \mathbf{a}_s \in 2^I\}$  being an arbitrary basis. In the box space  $\mathcal{R}$  each local box  $\{P(\mathbf{a}_s|\mathbf{s}) \mid \mathbf{s}, \mathbf{a}_s \in 2^I\}$  is represented by a vector

$$|P\rangle = \sum_{\mathbf{s} \in 2^I} \sum_{\mathbf{a}_s \in 2^I} P(\mathbf{a}_s|\mathbf{s}) |\mathbf{s}, \mathbf{a}_s\rangle = \int d\lambda \varrho_\lambda \bigotimes_{k \in I} \sum_{s_k=0}^1 |s_k\rangle_k \otimes |a_{k|s_k}^\lambda\rangle_k \quad (2)$$

that is a convex combination of  $4^n$  vectors in

$$P_n = \{|\mathbf{a}; \mathbf{b}\rangle = \bigotimes_{k \in I} (|0\rangle_k \otimes |a_k\rangle_k + |1\rangle_k \otimes |b_k\rangle_k) \mid \mathbf{a}, \mathbf{b} \in 2^I\}. \quad (3)$$

On the other hand every possible convex combination of vectors in  $P_n$  also corresponds to a possible set of statistical predictions by the local realistic model in the given scenario. Thus all possible statistical predictions of a local realistic model are identical with the vectors in the so-called Bell polytope  $\mathcal{B}_n = \text{Conv}(P_n)$  formed by all possible convex combinations of vectors in  $P_n$ . Those vectors in  $P_n$  are vertices, or extremal vectors, of the Bell polytope. Because of the following identities

$$|0; 0\rangle_k + |1; 1\rangle_k = |0; 1\rangle_k + |1; 0\rangle_k \quad (\forall k \in I), \quad (4)$$

where we have denoted  $|a; b\rangle = |0\rangle \otimes |a\rangle + |1\rangle \otimes |b\rangle$ , there are at most  $3^n$  linearly independent vectors in  $P_n$ . By noting the direct product structure of the vectors in  $P_n$ , it is straightforward to compute the Gramm matrix of all  $3^n$  vectors in  $P_n$  which can be readily found to be nonsingular. Thus the dimension can be exactly determined to be  $3^n$ . Therefore the linear span  $\mathcal{P}_n = \text{Span}(P_n)$  of  $P_n$ , i.e., the set of all vectors that are linear combinations of vectors in  $P_n$ , is of dimension  $\dim \mathcal{P}_n = 3^n$ . As a result the Bell polytope  $\mathcal{B}_n$  corresponding to  $(n, 2, 2)$  scenario is of dimension  $3^n - 1$ , i.e., there are  $3^n - 1$  independent parameters in a local box. Properties of Bell polytope in general scenario can be found in Ref.[24]. Those independent parameters in a local box can be taken as the following  $3^n - 1$  correlations

$$A_{\mathbf{s}}^c = \sum_{\mathbf{a}_s \in 2^I} (-1)^{\mathbf{a}_s \cdot \mathbf{c}} P(\mathbf{a}_s|\mathbf{s}) \quad (\mathbf{s} \subseteq c), \quad (5)$$

with properties  $A_{\mathbf{0}}^0 = 1$  and  $A_{\mathbf{s}}^c = A_{\mathbf{s} \wedge c}^c$  for a general  $\mathbf{s}$ .

In terms of Bell polytope, a Bell's inequality is nothing else than a hyper plane that divides the  $4^n$  dimensional box space  $\mathcal{R}$  into two parts with the Bell polytope being completely contained in one part, e.g.,  $\langle B|P\rangle \geq 0$  for all  $|P\rangle \in \mathcal{B}_n$  for a fixed vector  $|B\rangle \in \mathcal{R}$ . A facet of the Bell polytope is a hyper plane containing exactly  $3^n - 1$  linearly independent extremal vectors in  $P_n$  and naturally defines a Bell's inequality, which is referred to as a tight Bell's inequality. Precisely, a Bell's inequality  $\langle B|P\rangle \geq 0$

in the scenario  $(n, 2, 2)$  is called as tight if and only if there are  $3^n - 1$  independent extremal vectors  $|\mathbf{a}; \mathbf{b}\rangle \in P_n$  such that  $\langle B | \mathbf{a}; \mathbf{b} \rangle = 0$ . The following Lemma is about two standard forms of a Bell's inequality.

**Lemma** Every Bell's inequality can be cast into a form involving probabilities and a form involving correlations

$$\langle B | P \rangle = \sum_{\mathbf{s}, \mathbf{a}_s \in 2^I} B(\mathbf{a}_s, \mathbf{s}) P(\mathbf{a}_s | \mathbf{s}) = \sum_{\mathbf{c} \in 2^I} \sum_{\mathbf{s} \subseteq \mathbf{c}} B_s^c A_s^c \geq 0, \quad (6)$$

where two sets of coefficients are related to each other via

$$B_s^c = \frac{1}{2^{|\mathbf{c}|}} \sum_{\mathbf{a}_s \in 2^I} (-1)^{\mathbf{a}_s \cdot \mathbf{c}} B(\mathbf{a}_s, \mathbf{s}) \quad (7)$$

with real coefficients  $B(\mathbf{a}_s, \mathbf{s})$  satisfying

$$\sum_{\mathbf{s}, \mathbf{a}_s \in 2^I} B(\mathbf{a}_s, \mathbf{s}) = 1, \quad \sum_{\mathbf{a}_k | s_k = 0}^1 B(\mathbf{a}_s, \mathbf{s}) \text{ is independent of } s_k \quad (\forall k \in I). \quad (8)$$

**Proof.** Being determined by a hyper plane in  $\mathcal{R}$ , every Bell's inequality assumes the form  $\langle B | P \rangle \geq 0$  for some fixed vector  $|B\rangle \in \mathcal{R}$  and we have only to prove that its coefficients  $B(\mathbf{a}_s, \mathbf{s}) = \langle B | \mathbf{s} \rangle_I \otimes |\mathbf{a}_s\rangle_I$  can be chosen to satisfy conditions Eq.(8). For an arbitrary extremal vector  $|\mathbf{b}_0; \mathbf{b}_1\rangle \in P_n$  of the Bell polytope Bell's inequality must hold

$$\langle B | \mathbf{b}_0; \mathbf{b}_1 \rangle = \sum_{\mathbf{s} \in 2^I} B(\mathbf{b}_s, \mathbf{s}) \geq 0. \quad (9)$$

Since  $\{|P\rangle | \langle B | P \rangle = 0\}$  defines a hyper plane in the box space  $\mathcal{R}$  with one dimension less than the Bell polytope, there is at least one extremal vector  $|\mathbf{b}_0; \mathbf{b}_1\rangle \in P_n$  of the Bell polytope such that  $\langle B | \mathbf{b}_0; \mathbf{b}_1 \rangle \neq 0$ . That is to say the inequality Eq.(9) cannot be attained for all  $\mathbf{b}_s$  from which it follows

$$\langle B | \theta \rangle = \sum_{\mathbf{s}, \mathbf{b}_s \in 2^I} B(\mathbf{b}_s, \mathbf{s}) > 0, \quad |\theta\rangle = \sum_{\mathbf{s} \in 2^I} |\mathbf{s}\rangle_I \otimes \sum_{\mathbf{a} \in 2^I} |\mathbf{a}\rangle_I.$$

Thus the first condition in Eq.(8) can be satisfied by dividing both sides of Bell's inequality by a positive number  $\langle B | \theta \rangle$ . Secondly, we can rewrite Bell's inequality for probabilities in Eq.(6) as

$$0 \leq \sum_{\mathbf{c}, \mathbf{s} \in 2^I} \frac{B_s^c A_s^c}{2^{n-|\mathbf{c}|}} = \sum_{\mathbf{c} \in 2^I} \sum_{\mathbf{s} \subseteq \mathbf{c}} \frac{A_s^c}{2^{n-|\mathbf{c}|}} \sum_{\mathbf{t} \in 2^I} B_t^c \delta_{\mathbf{t} \wedge \mathbf{c}, \mathbf{s} \wedge \mathbf{c}} := \sum_{\mathbf{c} \in 2^I} \sum_{\mathbf{s} \subseteq \mathbf{c}} \tilde{B}_s^c A_s^c. \quad (10)$$

where we have used relation Eq.(7) and the fact  $A_s^c = A_{\mathbf{s} \wedge \mathbf{c}}^c$  (non-signaling). This new set of coefficients satisfy  $\tilde{B}_s^c = \tilde{B}_{\mathbf{s} \wedge \mathbf{c}}^c$  and consequently the coefficients

$$\tilde{B}(\mathbf{a}_s, \mathbf{s}) = \sum_{\mathbf{c} \in 2^I} \frac{(-1)^{\mathbf{a}_s \cdot \mathbf{c}}}{2^{n-|\mathbf{c}|}} \tilde{B}_{\mathbf{s} \wedge \mathbf{c}}^c \quad (11)$$

define the same Bell's inequality as  $B(\mathbf{a}_s, \mathbf{s})$  does while the second conditions in Eq.(8) are satisfied.  $\square$

Some remarks are in order. In the most general scenario probability and correlation forms of Bell's inequality can be obtained similarly, with coefficients  $B(\mathbf{a}_s, \mathbf{s})$  satisfying similar conditions as in Eq.(8), which are identical to the non-signaling conditions. Next we shall consider a special example of Bell's inequality.

### 3. Bell-Hardy's inequality

From the view point of the relation between the entanglement and nonlocality, the Bell-Hardy inequality can be regarded as a most natural generalization to  $n$  particles of the well known Bell-CHSH inequality. It was shown by Gisin that every pure entangled two-qubit state violate the Bell-CHSH inequality and recently it has been shown that Bell-Hardy's inequality can be used to established the equivalence between the nonlocality and entanglement for pure states. Bell-Hardy's inequality is based on Hardy's argument or test for nonlocality.

Let  $n$  space-like separated observers, labeled with  $I$ , perform two dichotomic measurements locally, with setting vector  $\mathbf{s}$  and corresponding outcome vector  $\mathbf{a}_\mathbf{s}$ . Let  $\mathbf{0}, \mathbf{1}$  denote the all zero or one vector and  $\mathbf{1}_j = [\delta_{jk}]_{k \in I}$ . For example a setting vector  $\mathbf{s} = \mathbf{0}$  together with an outcome vector  $\mathbf{a}_\mathbf{s} = \mathbf{1}$  means that observer  $k$  measures the observable  $A_{k|0}$  and obtains 1 as outcome for each  $k \in I$ . The probability of this event is denoted as  $P(\mathbf{1}|\mathbf{0})$ . Hardy's test consists of the following  $n + 2$  conditions

$$P(\mathbf{1}|\mathbf{0}) > 0, \quad P(\mathbf{0}|\mathbf{1}) = 0, \quad P(\mathbf{1}|\mathbf{1}_j) = 0 \quad (j \in I). \quad (12)$$

For any local realistic model, where the probabilities assume the local form Eq.(1), these conditions cannot be satisfied simultaneously. For example from the first condition it follows that there is a set of hidden variables  $\lambda$  with a nonzero measure such that the response functions  $a_{k|0}^\lambda = 1$  for every observer  $k \in I$ . For these hidden variables, it follows from last  $n$  conditions that  $a_{j|1}^\lambda = 0$  for all  $j \in I$ , which contradicts with the second condition. Hardy's test gives rise to the following Bell-Hardy inequality

$$\sum_{j \in I} P(\mathbf{1}|\mathbf{1}_j) + P(\mathbf{0}|\mathbf{1}) - P(\mathbf{1}|\mathbf{0}) \geq 0. \quad (13)$$

Hardy's test was for two particles and Mermin formulated Hardy's inequality for two qubits [17], which is equivalent to Bell-CHSH inequality and Cereceda [18] proposed the  $n$ -particle version. By using the notions of Bell polytope, we can rewrite Bell-Hardy inequality simply as  $\langle H_n | P \rangle \geq 0$  for all  $|P\rangle \in \mathcal{B}_n$  where

$$|H_n\rangle = \sum_{j=1}^n |\mathbf{1}_j\rangle_I \otimes |\mathbf{1}\rangle_I + |\mathbf{1}\rangle_I \otimes |\mathbf{0}\rangle_I - |\mathbf{0}\rangle_I \otimes |\mathbf{1}\rangle_I. \quad (14)$$

If a quantum mechanical state passes Hardy's test, i.e., all conditions in Eq.(12) are satisfied, then it naturally leads to a violation to Bell-Hardy inequality given above. On the other hand there are entangled pure states that fail Hardy's test while violating the Bell-Hard inequality. To prove that the Bell-Hardy inequality holds for any local realistic model we have only to show that for all extremal local boxes  $|\mathbf{a}; \mathbf{b}\rangle \in P_n$  the inequality holds

$$\begin{aligned} \langle H_n | \mathbf{a}; \mathbf{b} \rangle &= \sum_{j \in I} b_j \prod_{k \neq j} a_k + \prod_{k \in I} (1 - b_k) - \prod_{k \in I} a_k \\ &\geq \left( \sum_{j \in I} b_j + \prod_{k \in I} (1 - b_k) - 1 \right) \prod_{k \in I} a_k \geq 0. \end{aligned}$$

**Theorem 1** Bell-Hardy's inequality Eq.(13) is tight.

**Proof.** We need to show that there are  $3^n - 1$  independent extremal vectors  $|\mathbf{a}; \mathbf{b}\rangle \in P_n$ , defined in Eq.(3), that saturate Bell-Hardy inequality  $\langle H_n | \mathbf{a}; \mathbf{b} \rangle = 0$ . In order to count the dimension  $d_H = \dim \mathcal{V}_n$  of the subspace  $\mathcal{V}_n = \text{Span}(V_n)$  of all possible linear combinations of vectors in

$$V_n = \{|\mathbf{a}; \mathbf{b}\rangle | \langle H_n | \mathbf{a}; \mathbf{b} \rangle = 0\} \quad (15)$$

the basic idea is to add one more vector that does not belong to  $\mathcal{V}_n$ , namely  $|\mathbf{0}; \mathbf{0}\rangle$ , to  $V_n$  and prove that  $\mathcal{P}_n = \text{Span}(\tilde{V}_n)$  with  $\tilde{V}_n = \{|\mathbf{0}; \mathbf{0}\rangle\} \cup V_n$ , from which it follows  $d_H = \dim \mathcal{P}_n - 1$ . First of all we have  $|\mathbf{0}; \mathbf{0}\rangle \notin \mathcal{V}_n$  because  $\langle H_n | \mathbf{0}; \mathbf{0} \rangle = 1$  while all vectors in  $\mathcal{V}$  satisfy  $\langle H_n | V \rangle = 0$ . Secondly, we observe

$$\text{Span}(U_n \cup \{|\mathbf{0}; \mathbf{0}\rangle\}) = \text{Span}(U'_n \cup \{|\mathbf{0}; \mathbf{1}\rangle\}) = \mathcal{P}_n \quad (16)$$

where  $U_n = \{|\mathbf{a}; \mathbf{b}\rangle | \mathbf{b} \neq \mathbf{0}\}$  and  $U'_n = \{|\mathbf{a}; \mathbf{b}\rangle | \mathbf{a} \neq \mathbf{1}\}$ . To show this, let  $u_n$  be the number of independent vectors in  $U_n$  and obviously  $u_1 = 2$ . Denoting  $\mathbf{a} = (a_1, \mathbf{a}')$  and  $\mathbf{b} = (b_1, \mathbf{b}')$  with  $\mathbf{a}'$  and  $\mathbf{b}'$  being two  $(n-1)$  dimensional binary vectors, we have a partition  $U_n = \cup_{a_1, b_1=0,1} U_{a_1 b_1}$  where

$$U_{a_1 b_1} = \{ |a_1; b_1\rangle \otimes |\mathbf{a}'; \mathbf{b}'\rangle | (b_1, \mathbf{b}') \neq \mathbf{0} \} \quad (a_1, b_1 = 0, 1).$$

It is clear that  $U_{a_1,1} = |a_1; 1\rangle \otimes P_{n-1}$  has  $3^{n-1}$  independent vectors for each  $a_1 = 0, 1$  while  $U_{0,0} = |0; 0\rangle \otimes U_{n-1}$  has  $u_{n-1}$  independent vectors. All the vectors in  $U_{1,0}$  can be obtained by linear combinations of vectors in other 3 sets due to identity Eq.(4). Thus we have  $u_n = u_{n-1} + 2 \times 3^{n-1}$  which gives  $u_n = 3^n - 1$ . Moreover we have  $|\mathbf{0}; \mathbf{0}\rangle \notin \text{Span}(U_n)$ , due to the identity  $\langle \mathbf{a}; \mathbf{b} | \mathbf{1} \rangle_I = \langle \mathbf{b} |_I$ , and thus  $\text{Span}(U_n \cup \{|\mathbf{0}; \mathbf{0}\rangle\}) = \mathcal{P}_n$ . The statement for  $U'_n$  can be proved in the same manner by noting that  $|\mathbf{0}; \mathbf{1}\rangle \notin \text{Span}(U'_n)$ .

We shall proceed with induction. Obviously it is true if  $n = 1$ , i.e.,  $\mathcal{P}_1 = \text{Span}(\tilde{V}_1)$  with  $V_1 = \{|1; 1\rangle, |1; 0\rangle\}$ . Suppose  $\mathcal{P}_n = \text{Span}(\tilde{V}_n)$  is true for  $n$ . By denoting  $(\mathbf{a}, a_{n+1})$  and  $(\mathbf{b}, b_{n+1})$  two  $(n+1)$  dimensional binary vectors, it is easy to check the following statements are true

1.  $|\mathbf{a}, 0\rangle; (\mathbf{b}, 0)\rangle = |\mathbf{a}; \mathbf{b}\rangle \otimes |0; 0\rangle \in V_{n+1}$  if  $\mathbf{b} \neq \mathbf{0}$ , i.e.,  $U_n \otimes |0; 0\rangle \subset V_{n+1}$ , because

$$\langle H_{n+1} | \mathbf{a}, 0\rangle; (\mathbf{b}, 0)\rangle = \prod_{k \in I} \delta(b_k, 0);$$

2.  $|\mathbf{a}, 0\rangle; (\mathbf{b}, 1)\rangle = |\mathbf{a}; \mathbf{b}\rangle \otimes |0; 1\rangle \in V_{n+1}$  if  $\mathbf{a} \neq \mathbf{1}$ , i.e.,  $U'_n \otimes |0; 1\rangle \subset V_{n+1}$ , because

$$\langle H_{n+1} | \mathbf{a}, 0\rangle; (\mathbf{b}, 1)\rangle = \prod_{k \in I} \delta(a_k, 1);$$

3.  $|\mathbf{a}, 1\rangle; (\mathbf{b}, 0)\rangle = |\mathbf{a}; \mathbf{b}\rangle \otimes |1; 0\rangle \in V_{n+1}$  if  $|\mathbf{a}; \mathbf{b}\rangle \in V_n$ , i.e.,  $V_n \otimes |1; 0\rangle \subset V_{n+1}$ , because

$$\langle H_{n+1} | \mathbf{a}, 1\rangle; (\mathbf{b}, 0)\rangle = \langle H_n | \mathbf{a}; \mathbf{b}\rangle;$$

4.  $|\mathbf{0}; \mathbf{0}\rangle \otimes |1; 1\rangle \in V_{n+1};$

5.  $|\mathbf{1}; \mathbf{0}\rangle \otimes |1; 1\rangle \in V_{n+1}.$

By adding one single vector  $|\mathbf{0}; \mathbf{0}\rangle \otimes |0; 0\rangle \in \tilde{V}_{n+1}$  to the subset  $U_n \otimes |0; 0\rangle$  we see that  $\mathcal{P}_n \otimes |0; 0\rangle \subset \text{Span}(\tilde{V}_{n+1})$ , taking into account Eq.(16). In particular we have



$$6. |\mathbf{1}; \mathbf{0}\rangle \otimes |0; 0\rangle \in \text{Span}(\tilde{V}_{n+1}).$$

Because of observations 2 and 4, together with identity Eq.(4), we have  $|\mathbf{0}; \mathbf{0}\rangle \otimes |1; 0\rangle \in \text{Span}(\tilde{V}_{n+1})$  so that, due to observation 3,  $\mathcal{P}_n \otimes |1; 0\rangle = \text{Span}(\tilde{V}_n) \otimes |1; 0\rangle$  is also a subset of  $\text{Span}(\tilde{V}_{n+1})$ . In particular we have

$$7. |\mathbf{1}; \mathbf{0}\rangle \otimes |1; 0\rangle \in \text{Span}(\tilde{V}_{n+1}).$$

As a result of observations 5,6, and 7, taking into account identity Eq.(4), we have  $|\mathbf{1}; \mathbf{0}\rangle \otimes |0; 1\rangle \in \text{Span}(\tilde{V}_{n+1})$  from which it follows  $\mathcal{P}_n \otimes |0; 1\rangle$  is also a subset of  $\text{Span}(\tilde{V}_{n+1})$ , taking into account Eq.(16). All these results, together with identity Eq.(4), infer further  $\mathcal{P}_n \otimes |1; 1\rangle \subset \text{Span}(\tilde{V}_{n+1})$ , from which it follows  $\mathcal{P}_{n+1} = \cup_{x,y=0,1} \mathcal{P}_n \otimes |x; y\rangle \subseteq \text{Span}(\tilde{V}_{n+1})$ . On the other hand we have  $\tilde{V}_{n+1} \subseteq \mathcal{P}_{n+1}$ , meaning that  $\text{Span}(\tilde{V}_{n+1}) \subseteq \mathcal{P}_{n+1}$ , which leads to  $\text{Span}(\tilde{V}_{n+1}) = \mathcal{P}_{n+1}$  if  $\text{Span}(\tilde{V}_n) = \mathcal{P}_n$ .  $\square$

Similar methods have been used in [25] to derive some other tight Bell's inequalities.

#### 4. Non-signaling polytope

Quantum mechanical predictions violate Bell's inequality, rendering quantum theory nonlocal. However, quantum theory is not that nonlocal as it obeys the so called non-signaling conditions, i.e., local choices of measurement settings cannot be used to send information. More precisely, in the given  $(n, 2, 2)$  scenario, a set of  $4^n$  probabilities  $\tilde{P}(\mathbf{b}_t|\mathbf{t})$  are non-signaling if it holds

$$\sum_{b_k|t_k=0}^1 \tilde{P}(\mathbf{b}_t|\mathbf{t}) \text{ is independent of } t_k \quad (\forall k \in I). \quad (17)$$

That is to say, by summing out all the outcomes of the local measurement of a given observer, the joint probability of all other local measurements is independent of the choice of the measurement setting by the given observer. Obviously local boxes are non-signaling. From these non-signaling constraints, together with the normalization of the probability distribution for each given measurement setting, there are  $3^n - 1$  independent parameters in a general non-signaling joint probability. Conveniently these parameters can be taken as

$$\tilde{A}_t^c = \sum_{\mathbf{b}_t \in 2^I} (-1)^{c \cdot \mathbf{b}_t} \tilde{P}(\mathbf{b}_t|\mathbf{t}) \quad \text{for } \mathbf{t} \subseteq \mathbf{c} \quad (18)$$

with  $\tilde{A}_t^c = \tilde{A}_{t \cup c}^c$  and for a general  $\mathbf{t}$  reflecting the non-signaling constraints and  $\tilde{A}_0^0 = 1$  reflecting the normalization. Every set of non-signaling probabilities, referred to as a non-signaling box, also corresponds to a vector  $|\tilde{P}\rangle$  in the  $4^n$  dimensional box space  $\mathcal{R}$ . The non-negativeness conditions  $\tilde{P}(\mathbf{b}_t|\mathbf{t}) \geq 0$  for all possible  $4^n$  events define the non-signaling polytope  $\mathcal{P}_{ns}$  with facet given by conditions  $\tilde{P}(\mathbf{b}_t|\mathbf{t}) = 0$ . If a non-signaling box is uniquely determined by its zeros and normalization, then the facets defined by its zeros intersect at a single point, which is a vertex of the non-signaling polytope and it is called as an extremal box.



In certain sense quantum theory lies between the Bell polytope and the non-signaling polytope. On the one hand local realistic models cannot reproduce quantum theory and on the other hand there are non-signaling boxes, e.g., Popescu-Rohrlich box, whose correlations cannot be reproduced by quantum theory. Quite recently a thorough numerical searches for all the extremal tripartite boxes and tight Bell's inequalities has been performed and it turns out that they have the same number [26]. Later on this duality is rigorously established by Fritz [23]. Our second result is the following alternative explicit formula for this duality.

**Theorem 2** Every Bell's inequality in its standard form as given in Eq.(6) with  $B(\mathbf{a}_s, \mathbf{s})$  satisfying conditions Eq.(8) defines a non-signaling box

$$\tilde{P}(\mathbf{b}_t|\mathbf{t}) = \sum_{\mathbf{s} \in 2^I} B(\mathbf{b}_t + \mathbf{s} \wedge \mathbf{t}, \mathbf{s}), \quad (19)$$

where the addition of two binary vectors is always module 2. Every non-signaling box  $\{\tilde{P}(\mathbf{b}_t|\mathbf{t})|\mathbf{t}, \mathbf{b}_t \in 2^I\}$  defines the following Bell's inequality in its standard form

$$\sum_{\mathbf{s}, \mathbf{t} \in 2^I} \sum_{\mathbf{a}_s, \mathbf{b}_t \in 2^I} \frac{3^n \tilde{P}(\mathbf{b}_t|\mathbf{t}) P(\mathbf{a}_s|\mathbf{s})}{8^n (-3)^{|\mathbf{a}_s + \mathbf{b}_t + \mathbf{s} \wedge \mathbf{t}|}} = \sum_{\mathbf{c} \in 2^I} \sum_{\mathbf{s}, \mathbf{t} \subseteq \mathbf{c}} \frac{(-1)^{\mathbf{s} \cdot \mathbf{t}}}{2^{n+|\mathbf{c}|}} \tilde{A}_t^{\mathbf{c}} A_s^{\mathbf{c}} \geq 0. \quad (20)$$

If Bell's inequality Eq.(6) is tight then the corresponding non-signaling box Eq.(19) is extremal and if the nonlocal box  $\tilde{P}(\mathbf{b}_t|\mathbf{t})$  is extremal then the corresponding Bell's inequality Eq.(20) is tight.

**Proof.** Given an arbitrary Bell's inequality for probabilities as in Eq.(6), it is clear that  $\tilde{P}(\mathbf{b}_t|\mathbf{t})$  defined in Eq.(19) is nonnegative, due to the fact that

$$\tilde{P}(\mathbf{b}_t|\mathbf{t}) = \sum_{\mathbf{s}, \mathbf{a}_s \in 2^I} B(\mathbf{a}_s, \mathbf{s}) \delta(\mathbf{a}_s, \mathbf{b}_t + \mathbf{s} \wedge \mathbf{t}) = \langle B | \mathbf{b}_t; \mathbf{b}_t + \mathbf{t} \rangle \quad (21)$$

and  $|\mathbf{b}_t; \mathbf{b}_t + \mathbf{t}\rangle \in P_n$ . Here we have used the fact

$$\sum_{\mathbf{s}, \mathbf{a}_s} \delta(\mathbf{a}_s, \mathbf{b}_t + \mathbf{s} \wedge \mathbf{t}) |\mathbf{s}; \mathbf{a}_s\rangle = \bigotimes_{k \in I} (|0\rangle_k \otimes |b_{k|t_k}\rangle_k + |1\rangle_k \otimes |b_{k|t_k} + t_k\rangle_k) = |\mathbf{b}_t; \mathbf{b}_t + \mathbf{t}\rangle.$$

From conditions Eq.(8) it follows that Eq.(19) is a normalized probability distribution, since  $\langle B | \theta \rangle = 1$ , satisfying the non-signaling conditions, i.e.,  $\tilde{P}(\mathbf{b}_t|\mathbf{t})$  is a non-signaling box. If the given Bell's inequality is tight, i.e., there are  $3^n - 1$  independent extremal vectors such that  $\langle B | \mathbf{a}; \mathbf{b} \rangle = 0$  and we denote

$$V_B = \{|\mathbf{a}; \mathbf{b}\rangle \mid \langle B | \mathbf{a}; \mathbf{b} \rangle = 0\}.$$

For each extremal vector  $|\mathbf{a}; \mathbf{b}\rangle \in V_B$  we obtain a zero of the non-signaling box  $\tilde{P}(\mathbf{b}_t|\mathbf{t})$  by choosing the measurement settings according to  $\mathbf{t} = \mathbf{b} + \mathbf{a}$  with outcomes given by  $\mathbf{b}_t = \mathbf{a}$ , i.e.,  $\tilde{P}(\mathbf{a}|\mathbf{a} + \mathbf{b}) = 0$ . We note there exists a one-to-one correspondence

$$\begin{cases} \mathbf{a} = \mathbf{b}_t \\ \mathbf{b} = \mathbf{b}_t + \mathbf{t} \end{cases}, \quad \begin{cases} \mathbf{t} = \mathbf{a} + \mathbf{b} \\ \mathbf{b}_t = \mathbf{a} \end{cases}. \quad (22)$$

Because  $3^n - 1$  equations  $\langle B | \mathbf{a}; \mathbf{b} \rangle = 0$  plus the normalization  $\langle B | \theta \rangle = 1$  determine the vector  $|B\rangle$  uniquely, since  $|\theta\rangle \notin \text{Span}(V_B)$ , the corresponding  $3^n - 1$  zeros of  $\tilde{P}(\mathbf{b}_t|\mathbf{t})$

plus the normalization determine the non-signaling box uniquely, meaning that the non-signaling box is extremal, i.e., a vertex of the non-signaling polytope.

On the other hand, for a given non-signaling box  $\tilde{P}(\mathbf{b}_t|\mathbf{t})$  if we can solve from Eq.(19) some function  $B(\mathbf{a}_s, \mathbf{s})$ , then this function will define a Bell's inequality via Eq.(6) since  $\langle B|\mathbf{a}; \mathbf{b} \rangle = \tilde{P}(\mathbf{a}|\mathbf{a} + \mathbf{b}) \geq 0$  for all possible extremal vectors  $|\mathbf{a}; \mathbf{b} \rangle \in P_n$  of the Bell polytope. By denoting

$$B_s^c = \frac{1}{2^{|\mathbf{c}|}} \sum_{\mathbf{a}_s \in 2^I} (-1)^{\mathbf{a}_s \cdot \mathbf{c}} B(\mathbf{a}_s, \mathbf{s}), \quad B(\mathbf{a}_s, \mathbf{s}) = \sum_{\mathbf{c} \in 2^I} \frac{(-1)^{\mathbf{a}_s \cdot \mathbf{c}}}{2^{n-|\mathbf{c}|}} B_s^c$$

we obtain

$$\begin{aligned} \tilde{A}_t^c &= \sum_{\mathbf{b}_t \in 2^I} (-1)^{\mathbf{c} \cdot \mathbf{b}_t} \tilde{P}(\mathbf{b}_t|\mathbf{t}) \\ &= \sum_{\mathbf{b}_t \in 2^I} (-1)^{\mathbf{c} \cdot \mathbf{b}_t} \sum_{\mathbf{s} \in 2^I} B(\mathbf{b}_t + \mathbf{s} \wedge \mathbf{t}, \mathbf{s}) \\ &= \sum_{\mathbf{b}_t \in 2^I} (-1)^{\mathbf{c} \cdot \mathbf{b}_t} \sum_{\mathbf{s} \in 2^I} \sum_{\mathbf{c}' \in 2^I} \frac{(-1)^{(\mathbf{b}_t + \mathbf{s} \wedge \mathbf{t}) \cdot \mathbf{c}'}}{2^{n-|\mathbf{c}'|}} B_s^{c'} \\ &= 2^{|\mathbf{c}|} \sum_{\mathbf{s} \in 2^I} (-1)^{(\mathbf{s} \wedge \mathbf{c}) \cdot \mathbf{t}} B_s^c \\ &= 2^n \sum_{\mathbf{s} \subseteq \mathbf{c}} (-1)^{\mathbf{s} \cdot \mathbf{t}} B_s^c \end{aligned}$$

As a result we have

$$B_s^c = \frac{1}{2^{n+|\mathbf{c}|}} \sum_{\mathbf{t} \subseteq \mathbf{c}} (-1)^{\mathbf{s} \cdot \mathbf{t}} \tilde{A}_t^c, \quad (23)$$

from which the Bell's inequality for correlations follows, and

$$\begin{aligned} B(\mathbf{a}_s, \mathbf{s}) &= \sum_{\mathbf{c}, \mathbf{t}, \mathbf{b}_t \in 2^I} \frac{(-1)^{(\mathbf{a}_s + \mathbf{b}_t + \mathbf{s} \wedge \mathbf{t}) \cdot \mathbf{c}}}{2^{3n-|\mathbf{c}|}} \tilde{P}(\mathbf{b}_t|\mathbf{t}) \\ &= \sum_{\mathbf{t}, \mathbf{b}_t \in 2^I} \frac{\tilde{P}(\mathbf{b}_t|\mathbf{t})}{8^n} \prod_{k \in I} (1 + 2(-1)^{a_k|s_k + b_k|t_k + s_k t_k}) \\ &= \sum_{\mathbf{t}, \mathbf{b}_t \in 2^I} \frac{3^n \tilde{P}(\mathbf{b}_t|\mathbf{t})}{8^n (-3)^{|\mathbf{a}_s + \mathbf{b}_t + \mathbf{s} \wedge \mathbf{t}|}}, \end{aligned}$$

from which the Bell's inequality for probabilities follows. Again, for each zero of the non-signaling box  $\tilde{P}(\mathbf{b}_t|\mathbf{t}) = 0$  there is an extremal vector  $|\mathbf{b}_t; \mathbf{b}_t + \mathbf{t} \rangle \in P_n$  of Bell polytope such that  $\langle B|\mathbf{b}_t; \mathbf{b}_t + \mathbf{t} \rangle = 0$ , i.e.,  $|\mathbf{b}_t; \mathbf{b}_t + \mathbf{t} \rangle \in V_B$ . If the non-signaling box is extremal, i.e., uniquely determined by its zeros and normalization, then there are no less than  $3^n - 1$  extremal vectors of the Bell polytope in  $V_B$ . The conditions  $\langle B|\mathbf{a}; \mathbf{b} \rangle = 0$  with  $|\mathbf{a}; \mathbf{b} \rangle \in V_B$ , together with the normalization  $\langle B|\theta \rangle = 1$ , should determine the non-signaling box and thus  $|B\rangle$  uniquely, meaning that there are at least  $3^n - 1$  independent extremal vectors in  $V_B$ , i.e., the corresponding Bell's inequality is tight.  $\square$

As an example, let us consider the Bell-Hardy inequality Eq.(13) or equivalently  $\langle H_n | P \rangle \geq 0$  with  $|H_n\rangle$  given in Eq.(14). This form of Bell-Hardy inequality is not standard because the coefficients

$$\langle H_n | \mathbf{s} \rangle_I \otimes | \mathbf{a}_s \rangle_I = \sum_{j \in I} \delta_{\mathbf{s}, \mathbf{1}_j} \delta_{\mathbf{a}_s, \mathbf{1}} + \delta_{\mathbf{s}, \mathbf{1}} \delta_{\mathbf{a}_s, \mathbf{0}} - \delta_{\mathbf{s}, \mathbf{0}} \delta_{\mathbf{a}_s, \mathbf{1}} \quad (24)$$

do not satisfy the conditions in Eq.(8). First of all we have  $\langle H_n | \theta \rangle = n$  and by denoting

$$H_s^c = \sum_{\mathbf{a}_s \in 2^I} (-1)^{\mathbf{c} \cdot \mathbf{a}_s} \langle H_n | \mathbf{s} \rangle_I \otimes | \mathbf{a}_s \rangle_I = \sum_{j \in I} (-1)^{|\mathbf{c}|} \delta_{\mathbf{s}, \mathbf{1}_j} + \delta_{\mathbf{s}, \mathbf{1}} - (-1)^{|\mathbf{c}|} \delta_{\mathbf{s}, \mathbf{0}}$$

the Bell-Hardy inequality can be cast into the following correlation form

$$\begin{aligned} 0 \leq \langle H_n | P \rangle &= \frac{1}{2^n} \sum_{\mathbf{c}, \mathbf{s} \in 2^I} H_s^c A_s^c \\ &= \frac{1}{2^n} \sum_{\mathbf{c} \in 2^I} \sum_{\mathbf{s} \subseteq \mathbf{c}} A_s^c \sum_{\mathbf{t} \in 2^I} H_t^c \delta_{\mathbf{t} \wedge \mathbf{c}, \mathbf{s} \wedge \mathbf{c}} := n \sum_{\mathbf{c} \in 2^I} \sum_{\mathbf{s} \subseteq \mathbf{c}} A_s^c \tilde{H}_s^c \end{aligned}$$

where

$$\tilde{H}_s^c = \frac{1}{n 2^n} \left( (-1)^{|\mathbf{c}|} \sum_{j \in I} \delta_{\mathbf{s}, \mathbf{1}_j \wedge \mathbf{c}} + \delta_{\mathbf{s}, \mathbf{1} \wedge \mathbf{c}} - (-1)^{|\mathbf{c}|} \delta_{\mathbf{s}, \mathbf{0} \wedge \mathbf{c}} \right) \quad (25)$$

so that the following coefficients

$$\tilde{H}(\mathbf{a}_s, \mathbf{s}) = \sum_{\mathbf{c} \in 2^I} \frac{(-1)^{\mathbf{c} \cdot \mathbf{a}_s}}{2^{n-|\mathbf{c}|}} \tilde{H}_{\mathbf{s} \wedge \mathbf{c}}^c \quad (26)$$

give rise to the standard probability form of Bell-Hardy inequality. Thus the extremal non-signaling box corresponding to the Bell-Hardy inequality via Theorem 2 has the following correlations

$$\tilde{A}_t^c = 2^n \sum_{\mathbf{s} \subseteq \mathbf{c}} (-1)^{\mathbf{s} \cdot \mathbf{t}} \tilde{H}_s^c = \frac{\sum_{j \in I} (-1)^{\mathbf{t}_j \cdot \mathbf{c}} + (-1)^{\mathbf{t} \cdot \mathbf{c}} - (-1)^{|\mathbf{c}|}}{n} \quad (27)$$

where  $\mathbf{t}_j = \mathbf{1} + \mathbf{t} \wedge \mathbf{1}_j = [1 - t_j \delta_{kj}]_{k \in I}$  for each  $j \in I$ , from which we obtain finally the corresponding extremal non-signaling box, referred to as Hardy box here,

$$\tilde{P}_H(\mathbf{b}_t | \mathbf{t}) = \frac{\sum_{j \in I} \delta_{\mathbf{b}_t, \mathbf{t}_j} + \delta_{\mathbf{b}_t, \mathbf{t}} - \delta_{\mathbf{b}_t, \mathbf{1}}}{n}. \quad (28)$$

In the case of  $n = 3$  the Hardy box is the extreme box number 29 listed in [26]. We note that the single particle probability of obtaining outcome 0 or 1 is  $1/n$  or  $(n-1)/n$  for both two choices of local measurements. Moreover this extremal non-signaling box can be easily proved to be nonlocal, i.e., violating some Bell's inequalities, e.g., the Bell inequality due to Werner & Wolf [27] and Zukowski & Brukner [28].

## 5. Conclusions and discussions

We have shown that the Bell-Hardy inequality arising from Hardy's nonlocality test is tight and derive an extremal nonlocal box via an explicit duality between the Bell polytope and the non-signaling polytope in the  $(n, 2, 2)$  scenario due to Fritz. Most recently Hardy's test is generalized to the detection of genuine multipartite nonlocality [29, 30]. It has been shown that fully entangled symmetric pure states of  $n$  qubits as well as all fully entangled pure states are genuinely multipartite nonlocal. That is to say, to reproduce quantum correlations in a fully entangled symmetric state or a three particles pure states, those hybrid local/nonlocal realistic models have to be signaling. We believe that the extremal non-signaling box defined here is also genuinely multipartite nonlocal. Finally, as a possible generalization, the explicit duality between Bell polytope and non-signaling polytope worked out here may lead to a partial duality in other scenarios.

This work is funded by the Singapore Ministry of Education (partly through the Academic Research Fund Tier 3 MOE2012-T3-1-009).

## References

- [1] Bell J S 1964 *Physics* (Long Island City, N.Y.) **1** 195
- [2] Bell J S 1966 *Rev. Mod. Phys.* **38** 447
- [3] Brunner N, Cavalcanti D, Pironio S, Scarani V and Wehner S 2014 *Rev. Mod. Phys.* **86**, 419
- [4] Werner S F 1989 *Phys. Rev. A* **40** 4277
- [5] Gisin N, 1991 *Phys. Lett. A* **154** 201
- [6] Gisin N and Peres A 1992 *Phys. Lett. A* **162** 15
- [7] Clauser J F, Horne M A, Shimony A and Holt R A 1969 *Phys. Rev. Lett.* **23** 880
- [8] Chen J L, Wu C F, Kwek L C and Oh C H 2004 *Phys. Rev. Lett.* **93** 140407
- [9] Choudhary S K, Ghosh S, Kar G and Rahaman R 2010 *Phys. Rev. A* **81** 042107
- [10] Popescu S and Rohrlich D 1992 *Phys. Lett. A* **166** 293
- [11] Yu S X, Chen Q, Zhang C J, Lai C H and Oh C H 2012 *Phys. Rev. Lett.* **109** 120402
- [12] Greenberger D M, Horne M A and Zeilinger A, in *Bells Theorem, Quantum Theory, and Conceptions of the Universe*, edited by M. Kafatos (Kluwer, Dordrecht, 1989), p. 69; arXiv: 0712.0921.
- [13] Mermin N D 1990 *Phys. Rev. Lett.* **65** 3373
- [14] Tang W D, Yu S X and Oh C H 2012 *Phys. Rev. Lett.* **110** 100403
- [15] Hardy L 1992 *Phys. Rev. Lett.* **68** 2981
- [16] Hardy L 1993 *Phys. Rev. Lett.* **71** 1665
- [17] Mermin N D 1994 *Am. J. Phys.* **62** 880
- [18] Cereceda J L 2004 *Phys. Lett. A* **327** 433
- [19] Pawłowski M, Paterek T, Kaszlikowski D, Scarani V, Winter A and Żukowski M 2009 *Nature* **461** 1101
- [20] Cabello A 2013 *Phys. Rev. Lett.* **110** 060402
- [21] Fritz T, Sainz A B, Augusiak R, Brask J B, Chaves R, Leverrier A and Acin A 2013 *Nat. Comm.* **4** 2263
- [22] Popescu S and Rohrlich D 1994 *Found. Phys.* **24** 379
- [23] Fritz T 2012 *J. Math. Phys.* **53** 072202
- [24] Pironio S 2005 *J. Math. Phys.* **46** 062112

- [25] Augusiak R, Fritz T, Kotowski Ma, Kotowski Mi, Pawłowski M, Lewenstein M and Acín A 2012 *Phys. Rev. A* **85**, 042113
- [26] Pironio S, Bancal J D and Scarani V 2011 *J. Phys. A: Math. Theor.* **44** 065303
- [27] Werner R F and Wolf M M 2002 *Phys. Rev. A* **64** 032112
- [28] Żukowski M and Brukner Č 2002 *Phys. Rev. Lett.* **88** 210401
- [29] Chen Q, Yu S X, Zhang C J, Lai C H and Oh C H 2013 *Phys. Rev. Lett.* **112** 140404
- [30] Yu S X and Oh C H 2013 Tripartite entangled pure states are tripartite nonlocal *arXiv*: 1306.5330